Rare Variants in TP53 and Susceptibility to Neuroblastoma

Sharon J. Diskin, Mario Capasso, Maura Diamond, Derek A. Oldridge, Karina Conkrite, Kristopher R. Bosse, Mike R. Russell, Achille Iolascon, Hakon Hakonarson, Marcella Devoto, John M. Maris

Neuroblastoma is a cancer of the developing sympathetic nervous system that accounts for approximately 10% of all pediatric oncology deaths (1). Familial neuroblastoma is rare (approximately 1%), and most of these cases harbor germline mutations in ALK (2) or PHOX2B (3,4). Through a genome-wide association study of sporadic neuroblastoma, we have reported common single nucleotide polymorphisms (SNPs) within or upstream of CASC15 (5) and FLJ44180 (5), BARD1 (6), LMO1 (7), DUSP12 (8), HSD17B12 (8), DDX4/IL1RA4 (8), HACE1 (9), and LIN28B (9), along with a common copy number variation within NBPF23 (10), as each being highly associated with neuroblastoma. Collectively, however, these variants still account for only a small portion of the risk for developing neuroblastoma. Recent next-generation sequencing studies of paired tumor-normal genomes in neuroblastoma have revealed a striking paucity of somatic mutations (11–14); we hypothesized that rare variants in the host genome play an important role in tumorigenesis and explain a substantial proportion of the unidentified heritability in this disease.

Disruption or malfunction of the p53 pathway is a near-universal hallmark of cancer (15–17). Germline mutations in TP53 are the cause of Li-Fraumeni syndrome, and individuals who carry these mutations are at increased risk of developing a wide array of cancers at an early age (18). Recently, a rare polymorphism (rs78378222) in the 3′ UTR of TP53 was found to confer susceptibility to cutaneous basal cell carcinoma, prostate cancer, glioma, and colorectal adenomas, but not melanoma, colon cancer, or breast cancer (19). Somatic mutations of TP53 or other pathway members, such as MDM2, are rare in primary neuroblastomas obtained at diagnosis (14).

To identify germline variants associated with neuroblastoma at the TP53 locus, we performed genotype imputation on a discovery case series of 2101 neuroblastoma patients accrued and consented through the North American–based Children's Oncology Group and 4202 control subjects of European ancestry matched genetically and consented through the Children's Hospital of Philadelphia, as described previously (9). The Ethics Committee at the Children's Hospital of Philadelphia approved this study. All statistical tests were two-sided, and a P value less than 1.0 × 10⁻⁷ was considered statistically significant for subsequent replication efforts. Imputation was performed using IMPUTE2 (20) with default parameters and Ne equal to 20000, along with a multipopulation reference panel from the worldwide 1000 Genomes Project Phase 1 Interim release. After imputation, SNPs with minor allele frequency (MAF) less than 1% and/or IMPUTE2 info quality score less than 0.8 were removed. The remaining SNPs were tested for association with neuroblastoma using the frequentist association test under the additive model using the “score” method implemented in SNPTTEST (21). Although analysis of SNPs directly genotyped on the array revealed only modest evidence for association in the TP53 region (rs8079544: odds ratio [OR] = 1.3; 95% confidence interval [CI] = 1.2 to 1.6; P = 7.2 × 10⁻⁶) (Supplementary Table 1, available online), two imputed SNPs were highly associated with neuroblastoma (rs35850753: OR = 1.9, 95% CI = 1.5 to 2.3, P = 5.6 × 10⁻⁸; rs78378222: OR = 2.0, 95% CI = 1.6 to 2.7, P = 1.1 × 10⁻⁹) (Figure 1A). Imputation probabilities for rs35850753 and rs78378222 were very high (Supplementary Figure 1, available online).

To further confirm the accuracy of imputation, we performed polymerase chain reaction (PCR)–based genotyping on 176 case patients and compared these results to the most probable genotype based on imputation. We observed a 96% concordance rate overall (Supplementary Tables 2 and 3, available online).

The two neuroblastoma-associated SNPs map to UTRs of TP53 and are rare (MAF for rs35850753: 3.6% case patients, 1.9% control subjects; rs78378222: 2.7% case patients, 1.3% control subjects). The most statistically significant SNP, rs35850753, maps to the 5′ UTR of the Δ133 isoform of TP53, which is transcribed...
by an alternative promoter, and lacks a trans activation domain (Figure 1B). This isoform has been shown to have a dominant negative effect whereby it inhibits the tumor suppressive functions of full-length TP53. The other SNP, rs78378222, maps to the 3' UTR of TP53 and is the same rare variant found by Stacey and colleagues to confer susceptibility to several other cancers (19). It was further demonstrated that this rare allele disrupts the TP53 polyadenylation signal (AATA[A/C]A) resulting in impaired 3'-end processing of TP53 transcripts in blood and adipose tissue (19). We analyzed total RNA from two rs78378222 heterozygous primary neuroblastomas and observed a preference for the wild-type/protective allele in properly terminated and polyadenylated transcripts, whereas improperly terminated “run-on” transcripts were detected almost exclusively from the variant/risk allele (Supplementary Figure 2, available online). Based on data from the 1000 Genomes Phase I European population, rs35850753 and rs78378222 are in linkage disequilibrium ($r^2 = 0.52$ $D^' = 1$). Accordingly, rs35850753 was no

**Table 1.** Statistically significantly associated single nucleotide polymorphisms within TP53 at 17p13*

<table>
<thead>
<tr>
<th>SNP</th>
<th>A1/A2</th>
<th>Cohort†</th>
<th>Frequency A1 case patients</th>
<th>Frequency A1 control subjects</th>
<th>Pt</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs35850753</td>
<td>T/C</td>
<td>European ancestry</td>
<td>0.036 (n = 2101)</td>
<td>0.019 (n = 4202)</td>
<td>5.6 x 10^-9</td>
<td>1.9 (1.5 to 2.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>African ancestry</td>
<td>0.012 (n = 365)</td>
<td>0.004 (n = 2491)</td>
<td>1.3 x 10^-3</td>
<td>3.4 (1.5 to 7.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Italian</td>
<td>0.025 (n = 338)</td>
<td>0.005 (n = 781)</td>
<td>9.0 x 10^-5</td>
<td>4.7 (2.0 to 11.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Combined</td>
<td>0.026</td>
<td>0.012</td>
<td>3.4 x 10^-12</td>
<td>2.7 (2.0 to 3.6)</td>
</tr>
<tr>
<td>rs78378222</td>
<td>G/T</td>
<td>European ancestry</td>
<td>0.027 (n = 2101)</td>
<td>0.013 (n = 4202)</td>
<td>1.1 x 10^-8</td>
<td>2.0 (1.6 to 2.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>African ancestry</td>
<td>0.011 (n = 365)</td>
<td>0.002 (n = 2491)</td>
<td>4.2 x 10^-5</td>
<td>5.1 (2.0 to 12.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Italian</td>
<td>0.009 (n = 335)</td>
<td>0.001 (n = 753)</td>
<td>4.2 x 10^-2</td>
<td>4.6 (1.2 to 18.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Combined</td>
<td>2.0 x 10^-11</td>
<td>2.3 (1.8 to 2.9)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* A1, allele, on forward strand; A2: allele 2, on forward strand; CI, confidence interval; OR, odds ratio with respect to A1; SNP, single nucleotide polymorphism.
† No deviations from Hardy–Weinberg equilibrium were observed ($P > .001$) in all cohorts.
‡ Allelic $P$ values generated by SNPTEST score method [two-sided]; combined $P$ values from METAL using inverse-variance method [two-sided].
longer statistically significant after conditioning on the known functional SNP, rs78378222 ($P = 0.07$). Neither variant was associated with clinical or biological covariates (Supplementary Table 4, available online); however, given the low MAF of both SNPs, larger sample sizes are required before definitive conclusions can be drawn on subgroup association. Lastly, we tested the TP53 codon 72 Arg/Pro missense variant (rs1042522) for association with overall survival in 1809 neuroblastoma case patients; however, we were unable to replicate the recent report that codon 72 Pro/Pro predicts poor survival in neuroblastoma (23) (Supplementary Figure 3 and Supplementary Table 5, available online).

We next sought to replicate the rs35850753 and rs78378222 associations in an African ancestry cohort of 365 neuroblastoma case patients and 2491 genetically matched control subjects (9). We used IMPUTE2 (20) to infer genotypes in a 6-Kb region around the TP53 locus using data from the 1000 Genomes Project. Similar to the European ancestry samples, imputation probabilities were very high in individuals of African ancestry (Supplementary Figure 4, available online). Using the proportion of African admixture as a covariable to correct for varying degrees of admixture among our samples, we observed highly statistically significant associations at both SNPs (Table 1). Indeed, these were the two most statistically significant SNPs within the 6-Kb region surrounding TP53 (Supplementary Figure 5, available online).

Finally, we performed PCR-based genotype typing of rs35850753 and rs78378222 in an Italian cohort of 351 neuroblastoma case patients and 780 control subjects as another independent replication effort. Both SNPs showed statistically significant association in the same direction seen in the European and African ancestry samples (Table 1). Meta-analysis of 10 290 individuals from the three studies using the inverse-variance method within METAL (24) resulted in highly statistically significant associations with larger effect sizes than seen for common variants identified by genome-wide association studies (rs35850753: $OR = 2.7$, 95% CI 2.0 to 3.6, $P = 3.43 \times 10^{-12}$; rs78378222: $OR = 2.3$, 95% CI 1.8 to 2.9, $P = 2.03 \times 10^{-12}$) (Table 1).

We have used imputation to infer genotypes for variants in the 1000 Genomes Project not directly assayed on the SNP arrays used. It is possible that additional rare variants within or nearby TP53 are associated with neuroblastoma, but were not detected in our study because of limitations in the imputation process.

In conclusion, here we describe the first report of rare germline variation associated with neuroblastoma susceptibility. Our findings add to the complex repertoire of human cancers influenced by the TP53 network. Accurate knowledge of these rare variants within or nearby TP53 transcripts will enable future studies using whole-genome sequencing data to more accurately identify rare variants contributing to neuroblastoma tumorigenesis.

**References**


**Funding**

This work was supported in part by grants from the National Institutes of Health (R01-CA124709 to JMM; R00-CA151869 to SJD; P10-HD026979 to MD); the Giulio D’Angio Endowed Chair (to JMM); Andrew’s Army Foundation (to JMM); the Alex’s Lemonade Stand Foundation (to JMM); the Fondazione Italiana per la Lotta al Neuroblastoma (to MD); the Giulio D’Angio Endowed Chair (to JMM); and the Center for Applied Genomics at the Children’s Hospital of Philadelphia Research Institute (to HH).

**Notes**

S.J. Diskin and J.M. Maris designed the experiment. S.J. Diskin drafted the manuscript, performed SNP association study, and analyzed PCR-based genotyping and sequencing data. M. Devoto and S. J.
Diskin replicated SNP associations in the African ancestry cohort. M. Capasso replicated SNP associations in the Italian cohort. M. Diamond performed PCR-based genotype validation in the European and African ancestry cohorts. D. A. Oldridge assisted in evaluation of genotype imputation probabilities. K. R. Bosse assisted with biological interpretation of SNP associations. M. Diamond and M. R. Russell performed 3’ RACE experiments. K. Conkrite performed TP53 run-on experiments. H. Hakonarson generated and provided all control data for the European and African ancestry cohorts. All authors commented on or contributed to the manuscript.

The study sponsors had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.


Affiliations of authors: Division of Oncology (SJD, MDi, DAO, KC, KRB, MRR, JMM), Center for Childhood Cancer Research (SJD, MDi, DAO, KC, KRB, MRR, JMM), Center for Applied Genomics (HH), and Division of Genetics (HH, MDe) Children’s Hospital of Philadelphia, Philadelphia, PA; Department of Pediatrics (SJD, DAO, KRB, HH, MDe, JMM), Abramson Cancer Center (SJD, JMM), Genomics and Computational Biology, Biomedical Graduate Studies (SJD, DAO, JMM), and Department of Biostatistics and Epidemiology (MDe), Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA (SJD, HH, MDe, JMM); Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università degli Studi di Napoli Federico II, Naples, Italy (MC, AI); Ceinge–Biotecnologie Avanzate, Naples, Italy (MC, AI); University of Rome “La Sapienza,” Department of Molecular Medicine, Rome, Italy (MDe).